VISIT…

LANZAROTE Caliente.COM

Материалы XI Всероссийского семинара "Нейроинформатика и ее приложения", Красноярск, 2003. - 215с. - С.176-177.

К ОПРЕДЕЛЕНИЮ ИНФОРМАТИВНОСТИ НЕЗАВИСИМЫХ ПЕРЕМЕННЫХ ДЛЯ НЕЙРОННОЙ СЕТИ

Царегородцев В.Г.

www.NeuroPro.ru tsar@neuropro.ru

Причиной написания заметки стало непрекращающаяся неправильная семантическая интерпретация пользователями информации о значимости входных сигналов нейронной сети. А именно, использование выданных программой-нейроимитатором значений показателей значимости для принятия решения об информативности или неинформативности входного сигнала.

В красноярской школе нейроинформатики под показателем значимости традиционно понимается и используется "оценка изменения значения целевой функции при заданном изменении параметра, взятая в первом приближении" [1]. Т.е. используется тейлоровское разложение в ряд – показателем значимости служит значение первой производной целевой функции по заданному параметру, домноженное на планируемое изменение параметра. При рассмотрении ситуации применительно к единственному обрабатываемому примеру выборки сразу возникают такие возражения:

1. "Оценка изменения значения целевой функции". Поскольку часто используется целевая функция с допуском на точность решения (для задач нейрорегрессии см. библиографию в [2], для нейроклассификации – целевая функция вида "расстояния до множества правильно интерпретируемых ответов" [1] при задаваемом уровне уверенности), то нельзя взять ту величину допуска, с которой нейросеть обучалась – значения целевой функции, производной и в итоге показателя значимости будут равны нулю. Поэтому допуск на время рассчета значимости ужесточается, что некорректно, так как характеризует не ту ситуацию, которой обучалась нейросеть.
2. "Оценка, взятая в первом приближении". Поскольку не оцениваются слагаемые второго и последующих порядков, не гарантируется, что приближение первого порядка в некотором конкретном случае будет давать возможность пренебрегать последующими компонентами разложения.
3. "Оценка, взятая в первом приближении". Поскольку при вычислении производных, выполняемом методом обратного распространения, участвуют и адаптивные параметры (веса синапсов) сети, то разные нейронные сети будут давать разные значения показателей значимости, вплоть до прямо противоположных картин.
4. "Домножение на планируемое изменение параметра". Под изменением входного сигнала понимается его исключение из рассмотрения, т.е. замена нулем на входе сети. Нуль чаще всего не будет соответствовать значению среднего для отмасштабированного (перед подачей сети, для приведения диапазона изменения исходного сигнала в диапазон, воспринимаемый нейросетью) сигнала, т.е. в этом случае статистически оправданный вариант замены не гарантируется.

Для некоторой выборки затем суммируют показатели значимости её примеров. Причем суммирование может вестить так, чтобы дать интересующий в данный момент результат: можно либо сделать акцент на высокой значимости входного сигнала для малого числа примеров, либо, наоборот, пренебречь отдельными экстремальными чувствительностями ради картины "в среднем". В [1] дано 2 способа усреднения показателей значимости по выборке, в авторской программе-нейроимитаторе NeuroPro на настоящий момент реализовано 4 таких способа. Каждый из способов даст свои значения показателей значимости и ранжирование входных сигналов в соответствии с ними. Вдобавок, опять возникают возражения непосредственно по базовой идее:

5. Вычисление производной происходит при имеющихся в обучающей выборке значениях других входных сигналов. Если обучающая выборка не информативна, то

распределение значений входных сигналов, которые в будущем будут поступать на нейросеть, будет отличаться от исходного, делая построенную из статистических соображений (усреднение, т.е. учет в соответствии с плотностью распределения) картину значений показателей значимости и их интерпретацию неадекватной.

6. Для скоррелированных входных сигналов их показатели значимости могут быть любыми в границах предельных случаев: "значимость у одного сигнала высока, у других низка, т.е. один входной сигнал стягивает одеяло на себя" – "значимости, т.е. вклады сигналов, примерно одинаковы". Т.о., нужен как предварительный анализ линейных и нелинейных корреляций между входными сигналами, так и попытка на основании рассмотрения результатов вычисления показателей значимости несколькими сетями определить реальное наличие групп взаимосвязанных показателей (поскольку нейросети могут, допустим, всегда выделять сигнал с наименьшим уровнем аддитивного шума и, таким образом, не формировать группу взаимосвязанных значимостей в полном соответствии с группировкой на основании корреляций) и степень вариативности показателей значимости внутри таких групп.

Высказанные возражения во многом известны (см., например, [3]). Можно сказать, что показатели значимости в описанном виде $\left\langle \partial H / \partial x \cdot (x - x^0) \right\rangle$, где $\langle \bullet \rangle$ означает усреднение в той или иной норме, реально могут использоваться только для задач автоматического сокращения числа входных сигналов нейронной сети (когда погрешности при определении значимости допустимы, поскольку чаще всего не слишком фатальны, и нейросеть после удаления даже реально не самого незначимого входного сигнала всегда дообучается до нужных требований к точности решения на выборке). Субъективное же нагружение смыслом значений-результатов вычисления по той или иной формуле рассчета значимости недопустимо, поскольку дает старт дальнейшим методологическим ошибкам (субъективное проведение границы при разделении на значимые и незначимые признаки и т.д.).

Автор отказался от использования количественных значений показателей значимости и действует на качественном уровне, путем обучения большого числа нейросетей и последующего сокращения числа их входных сигналов получая набор сетей, решающих задачу с нужной точностью на меньших, по сравнению с исходным, наборах входных сигналов. Те сигналы, которые оказались исключенными у всех или подавляющего большинства сетей, понимаются как неинформативные. Сигналы, используемые подавляющим большинством сетей, считаются информативными и обязательными для решения задачи с заданной точностью. Остальные трактуются как взаимозаменяемые сигналы и группы сигналов. Результаты применения указанной схемы даны в [4]. Расширение подхода для явного формирования тех или иных групп дублирующих признаков может строиться на алгоритмах [5].

Литература

1. *Горбань А.Н., Россиев Д.А.* Нейронные сети на персональном компьютере. Новосибирск: Наука, 1996. - 276с.
2. *Царегородцев В.Г.* Робастная целевая функция для задач нейроаппроксимации / Материалы настоящего семинара.
3. *Sarle W.S.* How to measure importance of inputs? SAS Institute Inc., Cary, NC, USA, 1999. ftp://ftp.sas.com/pub/neural/importance.html
4. *Царегородцев В.Г., Назимова Д.И.* Нейросетевое исследование значимости климатических параметров для идентификации зональных групп лесообразователей Сибири / Материалы VII Всеросс. семинара "Нейроинформатика и ее приложения", Красноярск, 1999. Красноярск.: КГТУ, 1999. - 167с. – С.152-153.
5. *Горбань А.Н., Миркес Е.М., Вашко Т.А.* Алгоритмы поиска дублирующих признаков. ИВМ СО РАН, Красноярск, 2000. – 42с. – депонирована в ВИНИТИ 24.05.00, №1501-800.